IMPROVING COGNITIVE INDEPENDENCE OF DEMENTIA PATIENTS BY DIRECTING THEM TO THE APPROPRIATE MUSIC THERAPY SESSIONS WHILE ANALYZING THEIR EMOTIONAL STATE

Project Id: TMP-2023-081

Hiththatiyage D.K.
IT20162696

B.Sc. (Hons) Degree in Information Technology
(Specialization in Software Engineering)

Department of Computer Science and Software Engineering
Sri Lanka Institute of Information Technology
Sri Lanka

September 2023

IMPROVING COGNITIVE INDEPENDENCE OF DEMENTIA PATIENTS BY DIRECTING THEM TO THE APPROPRIATE MUSIC THERAPY SESSIONS WHILE ANALYZING THEIR EMOTIONAL STATE

Project Id: TMP-2023-081

Hiththatiyage D.K.
IT20162696

The dissertation was submitted in partial fulfilment of the requirements
for the B.Sc. Special Honors degree in Information Technology (Specialization in Software Engineering)

Department of Computer Science and Software Engineering
Sri Lanka Institute of Information Technology
Sri Lanka

September 2023

DECLARATION

I declare that this is my own work, and this thesis does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature
Hiththatiyage D.K.	IT20162696	

The above candidate is carrying out research for the undergraduate Dissertation under my supervision.

Signature of the supervisor   Date

………………………….   ……………..…………

(Mrs Geethanjali Wimalaratne)

ABSTRACT

Dementia patients are individuals who have a decline in their cognitive and memory abilities that interferes with their daily life and activities. It is a progressive condition that affects thinking, memory, behavior, and the ability to perform everyday activities. Even though finding a cure for dementia like debilitating condition is a priority many researchers have now identified the need to provide a better quality of life and care for these patients. Most of the research’s state that these patients are in need of a third-party care to carry on with their day-to-day life. This might be a hectic procedure for the patient as well as these third-party members. So, in order to overcome this, we are implementing a solution to improve the quality of life and independence of these patients. Here we mainly consider the mild and the moderate party of this patients. People with dementia lose awareness of their emotional state, which changes periodically without the patient's knowledge. Therefore, the majority of experts in the sector advise musical therapy to keep patients emotional states stable. It has been discovered that a person's personality and mood swings are closely tied to music. And the part of the brain that affects emotional stability is where the timber, pitch, and meter are managed. Keeping the aforementioned considerations in mind, as well as the need to improve these patients' quality of life, this research study will discuss the implementation of an emotional-based personalized music player designed with dementia patients in mind.

Keywords: Dementia, emotional-based, Music player, cognitive, mild, moderate.

ACKNOWLEDGEMENT

First and foremost, I would like to express my sincere gratitude to my supervisor, Mrs. Geethanjali Wimalaratne, for all of her encouragement, counsel, and support during the research procedure as well as for the valuable insights she offered to ensure the research project's success. Along with my supervisor, Ms. Wishalya Tissera is the co-supervisor of this research project, and I'd want to thank her for always being helpful. This research required the perseverance of many people. Over the course of my educational endeavor, I was greatly helped by their priceless advice and time. I also like to thank the CDAP staff for helping us finish our project effectively. Then, I'd like to thank all of my friends from the bottom of my heart for their unwavering support throughout this journey. In addition, I'd want to thank my parents for their ongoing encouragement and support. A special thank you to the other researchers that collaborated with me to improve and further this research.

DECLARATION i

ABSTRACT ii

ACKNOWLEDGEMENT iii

LIST OF FIGURES vi

LIST OF TABLES vii

LIST OF ABBREVIATIONS viii

1. INTRODUCTION 1

1.1. Background & Literature Survey 2

1.2. Research Gap 5

1.3. Research Problem 10

2. RESEARCH OBJECTIVES 12

2.1. Main Objective 12

2.2. Specific Objectives 12

3. METHODOLOGY 14

3.1. Methodology 14

3.1.1. System Architecture 15

3.1.2. Data collection methods 16

3.2. Commercialization aspects of the product 17

3.3. Testing and Implementation 19

3.3.1. Implementation 19

3.3.2. Testing 25

4. RESULTS AND DISCUSSION 28

4.1. Results 28

4.2. Research Findings 30

4.3. Discussion 31

4.4. Summary of student contribution 32

5. CONCLUSION 33

REFERENCES 34

APPENDICES 36

Appendix A: Online Survey 36

Appendix B: Gantt chart 42

Appendix C: Work Breakdown Structure 42

Appendix D: Plagiarism report 43

LIST OF FIGURES

Figure 1.1.1 - Survey responses on the importance of a digital assistant 3
Figure 1.1.2 - Survey results on the effectiveness of a third party....................................4
Figure 1.1.3 - Survey results on the suggestions for dementia individuals to introduce music 4
Figure 1.1.4 - Survey results on accuracy of the available tools.......................................5
Figure 1.2.1 - Survey results on flaws in available tools 8
Figure 1.2.2 Survey results on the suggestions for a new assistant tool 8
Figure 1.3.1 Survey results on the peoples knowledge to the approach on music.10
Figure 1.3.2 Survey results on challenges of forecasting the emotional state 11
Figure 3.1.1 - Overall System Diagram 15
Figure 3.1.2 - System Diagram for the Individual Component 16
Figure 3.3.1 – Usage of the CNN model...20
Figure 3.3.2 – Training of the proposed model...21
Figure 3.3.3 - Getting the person age and the predicted emotion 23
Figure 4.1.1 - Accuracy of the generated attributes...28
Figure 4.1.2 - Sample image tested with a moderate dementia patient and the detected emotion..29
Figure 4.1.3 – Sample image of a mild dementia individual and the result.....................29
Figure 4.1.4 - Sample image tested with a team member 30

LIST OF TABLES

Table 1.2.1 - Research gap compared to existing systems...9
Table 3.3.1 -External Tools..17
Table 3.3.2 - Test cases for Emotion generation and music player 25
Table 3.3.3 - Test cases done by end users 27
Table 3.3.4 - Test cases for devises with different OS...25
Table 3.3. - Test cases for devises with different OS...25

LIST OF ABBREVIATIONS

Abbreviation	Description
CNN	Convolutional Neural Network
UI	User Interface
DL	Deep Learning
API	Application Programming Interface
URL	Uniform Resource Locater

1. INTRODUCTION

Dementia patients are individuals who have a decline in their cognitive and memory abilities that interferes with their daily life and activities. It is a progressive condition that affects thinking, memory, behavior, and the ability to perform everyday activities. The number of elderly persons and life expectancy are both rising as a result of developments in medicine and technology. It is anticipated that every 20 years, the number of people living with Alzheimer's disease will nearly double [1]. The family has to take on the role of primary caregiver, which is frequently an emotionally challenging job. Family caregivers made up around one-third of those who showed symptoms of depression. Initially, it can be quite difficult for distant caregivers to continuously monitor patients as not everyone is living closer by [2]. It has been demonstrated that engaging in activities improves quality of life for those who have dementia, some activities may become impossible due to the disease's symptoms. In most circumstances their only option is to depend on a third party to carry on with these activities. However, finding accessible activities and devices to improve the quality of life of these patients can be difficult for family members and caregivers, which will be examined more in this study. Usage of touch displays has been found to be successful for dementia patients and is gaining attraction. There is a physical link between the user and the display, and when individuals touch the screen, they immediately receive feedback [3]. One of the most important and difficult tasks in developing such applications is creating a UI that is suitable for individuals who are suffering from Dementia. Today, mobile devices have become the standards for the implementation of assistive technologies to help people with physical and cognitive disabilities. The user experience for these apps has to be straightforward and uncomplicated, with larger fonts and clearly labeled, large buttons. The GUI needs to be adjusted for older users with poor vision. It shouldn't rely on the user having solid motor abilities as well. To the greatest extent possible, colors should be employed to further separate the various purposes of a button or region [4]. Recent studies indicate, people with dementia can experience and enjoy music even in their latter stages of the disease [3]. To enhance cognitive performance, there are two methods: pharmaceutical and non-pharmacological interventions. People with dementia still preserve their musical ability. Hence, one of the main strategies of the non-pharmacological intervention approach is music-based intervention [4]. Even though there are benefits associated with offering musical activities, caregivers lack the tools and sometimes the expertise to assist the patient in selecting music in accordance with preference or mood. So here mainly in this study the author will examine on a method to induce these patients with music.

1.1. Background & Literature Survey

There are several ways to extract face and audio elements from an audio signal, but very few of the systems created can generate an emotion-based music playlist based on emotions. This component's main goal is to improve the previous system's weaknesses by creating an automatic emotion-based music generator that creates a personalized playlist using user-extractable facial features. According to the findings from past studies, we can see that this component of the research has been researched by many. Various tools have been built to aid dementia patients in many ways. However, majority of them are built up for the use of the caregiver and they are missing some critical factors that should be improved to be used by the patient themself. According to the referred information, an assistant is mandatory for these individuals to regularly engage with the rest of the people and to improve their quality of life. As shown in Figure 1.1-1 which denotes the data retrieved by the survey implemented, it is validated that most of the people stated that a digital assistant is mandatory for these individuals. According to the referred information, a digital assistant is mandatory for cognitive disabled individuals to improve their quality of life and independence with the help of music (Figure 1.1). This assistant will guide these people to necessary music sessions in accordance with those emotions.

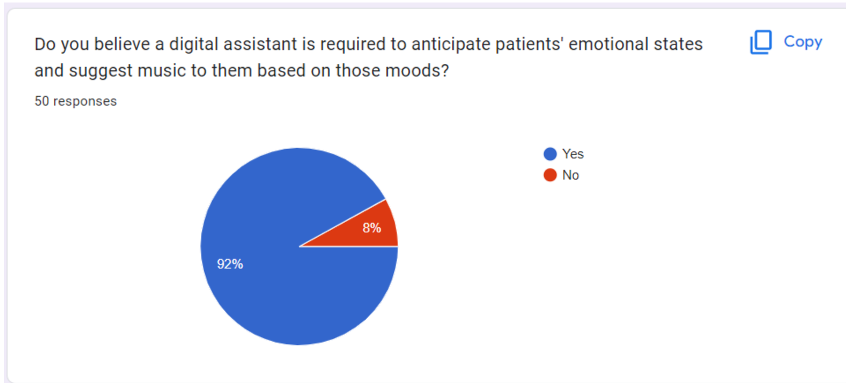


Figure 1.1.1 - Survey responses on the importance of a digital assistant

However, as already mentioned, the main method by which these dementia patients are directed towards musical therapy sessions is with the assistance of a third party. Given their current emotional condition, these people cannot be made open to more precise or successful therapy sessions, Figure 1.1-2 depicts it too. Some people have remarked that it is not always effective to direct dementia patients to necessary music therapy sessions when a third party is involved. They have stated that the current mood these individuals are in is particularly important when it comes to these sessions, and these caregivers/family members will not always be able to discern it accurately and these people to help with cannot be around all the time. So, it emphasizes that, a smart assistant is very important for a dementia individual when it comes to conducting music therapy sessions.

According to the data gathered by the survey, the pie chart in Figure 1.1-3 shows that the majority of respondents believe using an assistive tool makes it easier for a person with dementia to be directed to the right music therapy sessions according to their emotions to improve their quality of life and independence, while the least amount of respondents' support receiving assistance from a third party.

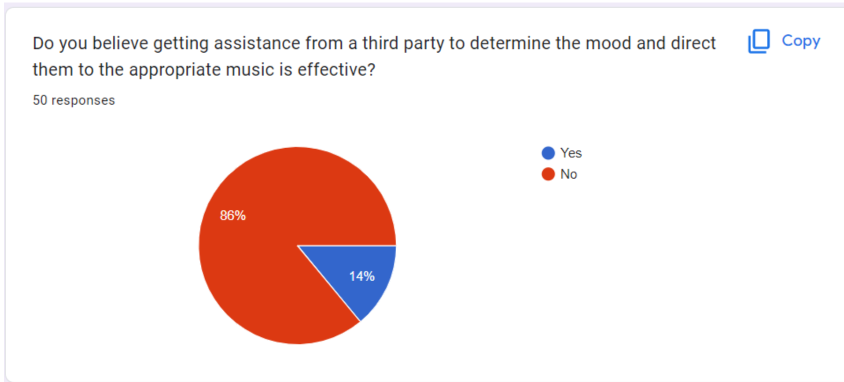


Figure 1.1.2 - Survey results on the effectiveness of a third party

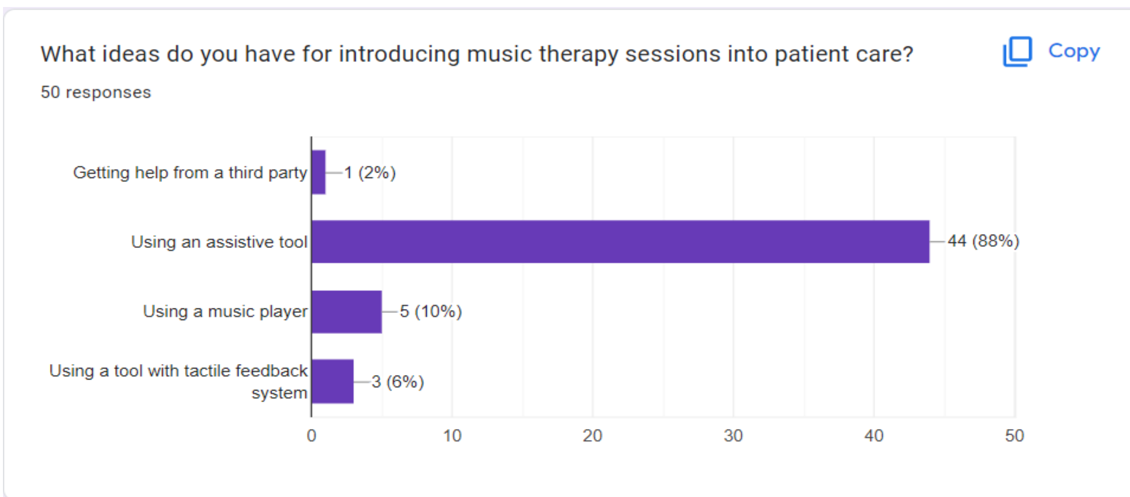


Figure 1.1.3 - Survey results on the suggestions for dementia individuals to introduce music

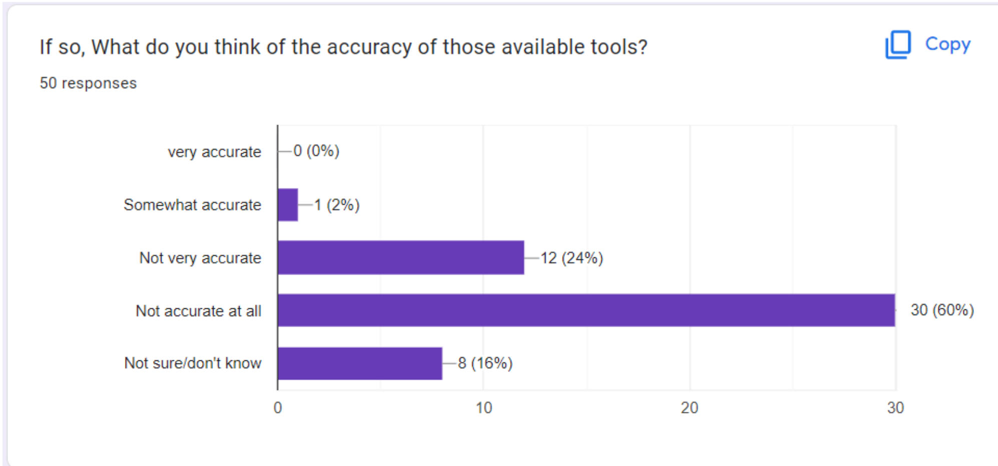


Figure 1.1.4 - Survey results on accuracy of the available tools

Moreover, technical improvements have made it possible to recognize emotions and predict music in accordance with those using a variety of software and apps. As seen in Figure 1.1-4 mandatory responses are for low accuracy for existing digital assistants. Many claims that the devices currently available do not accurately detect emotions and are not user-friendly for dementia patients. The literature reviews in article [5] vividly illustrate the operation of similar systems that are already in place, as well as their benefits and drawbacks.

1.2. Research Gap

There are several ways to separate face and audio components from an audio stream, but relatively few algorithms have been developed that can produce a music playlist based on emotions expressed by people. The primary objective of this component is to strengthen the shortcomings of the previous system by developing an automatic emotion-based music generator that generates customized playlists using user-extractable facial data.

Hafeez Kabani's research project A [5] centered on developing a music player based on human emotions, where the user's image is either captured via a webcam or can be accessed from the stored image on the hard drive. One limitation of this research is that it primarily focuses on Windows programs rather than mobile applications, which will be a more approachable method for addressing individuals. The suggested technique also has a tendency to deliver unpredictable outcomes under very poor camera quality and really poor lighting conditions and is not carried out with a dementia audience in mind. As a result, it is also not user-friendly enough for someone with dementia.

A system called Memory Tracks, an android application that uses music related with daily duties, has been proposed by Stuart Cunningham [6] (Research B) for Dementia patients. With the benefit of song-task association, this program seeks to support daily routines, help with care, help with agitation management, and help trigger memory. Here, instead of considering the patient's emotional state, the appropriate music was selected for each activity from the song library and matched to each resident using their demographic information, such as their birth year and where they spent their childhood.

Based on the documentary Alive Inside, which explores the impact of music on older people's subconscious, Alive Inside [7] (Research C) is a customized music-streaming app for dementia patients. It stands out from other music apps since it gives users a personalized music listening list based on their unique lives to enhance their quality of life. One of this application's biggest flaws is that when creating the user's personalized profile, the developers have not taken their emotional condition into account.

Study D [8] uses Real-time EEG to recognize emotions in music therapy. Here, hardware such as the PET 2 and Emotiv wireless headset were used for gathering the EEG data. Human computer interfaces may take on a new dimension with real-time EEG-enabled interaction. But, adding further hardware, such as sensors or EEG, tends toraise the price of the suggested design, which can be cited as a drawback.

Pranjul Agrahari came up with a system for musical therapy using facial expressions [9] (Research E) where the features are captured from a web camera and generates a playlist of relevant songs for the patient. One of the major drawbacks of this project is that here a small dataset of songs from a local collection is used which makes the system less accurate. Similar to research A, this method only focuses on a computer-based system and tends to produce unpredictable results in low-quality lighting and camera settings.

As previously noted, numerous methods have been used to anticipate music in accordance with emotional state, however they have significant limitations, such as

I. For extracting facial features in real time, current methods are quite sophisticated in terms of time and memory needs.
II. Existing methods are less accurate at creating playlists based on a user's current emotional state and behavior and they have failed to produce a unique playlist according to the emotions and the age of the individual.
III. Several current systems need additional hardware to generate an automated playlist, which raises the overall cost.
IV. Several systems now in use produce unpredictable results when subjected to lighting and camera conditions that are both quite poor.
V. These systems do not focus on keeping the patient engaged throughout these sessions.

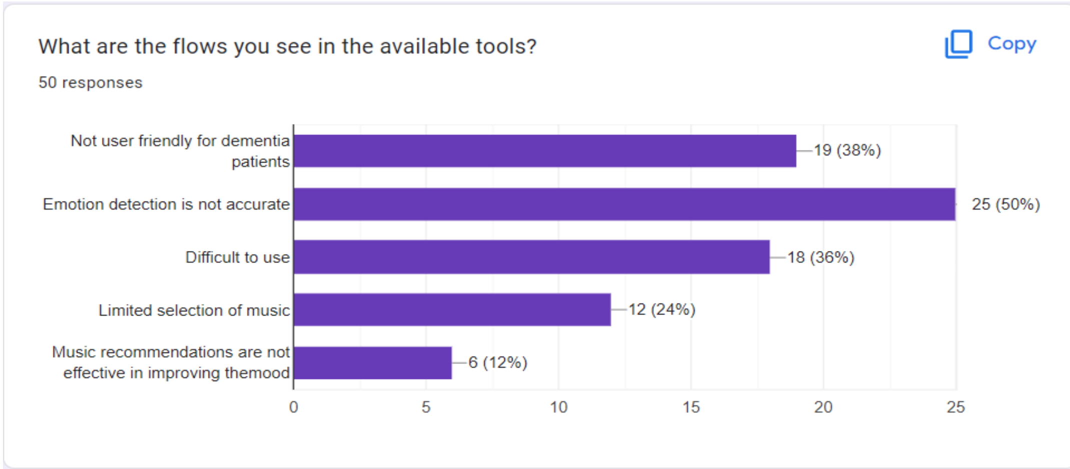


Figure 1.2.1 - Survey results on flaws in available tools

Also, the majority of survey participants claimed that the current methods do not accurately identify emotions and are difficult for people with dementia to use (Figure 1.2-1). In consideration of Figure1.2-2, many people suggest that when creating the new digital assistant, increasing the effectiveness of emotion recognition, offering a wider variety of music, using simple English and UIs will be beneficial.

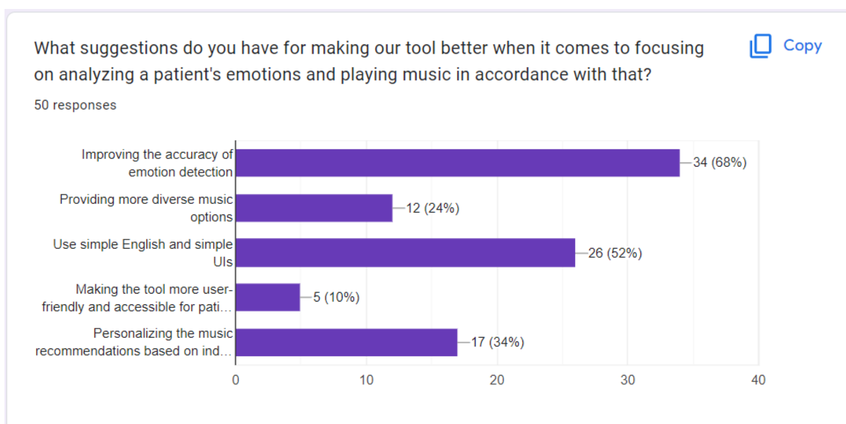


Figure 1.2.2 Survey results on the suggestions for a new assistant tool

Table 1.2.1 Research gap compared to existing systems

	Research A	Research B	Research C	Research D	Research E	Our Solution
Assist Emotion FeatureExtraction	Yes	No	No	Yes	Yes	Yes
optimized for mobile/ cloud use	No	Yes	Yes	No	No	Yes
Detect Age and classify music separately	No	Yes	Yes	No	No	Yes
Visual Presentations accordancewith the classified Music	No	No	No	No	No	Yes
Build A music Library based onreal time patient Reaction	No	No	No	No	No	Yes

Table 1.2-1 also makes a brief comparison between the suggested solution and the identified problems with the current systems. Reviewing the results reveals that this solution is implemented with far more innovative functionalities than other studies that are currently being done.

1.3. Research Problem

According to the survey conducted, it concludes that it's challenging to forecast these patients' emotional states because they change so frequently. Paper [10] also depicts that dementia increasingly affects memory, reasoning, language, and daily functioning in affected individuals. Dementia frequently accompanies emotional and behavioral issues and can lower a person's quality of life. People with dementia may find it challenging to express themselves verbally as the disease progresses, but even when they are unable to speak, they may still be able to hum or move to the music. Hence, music therapy may be especially beneficial for those who have dementia. Emotional responses frequently vary for people with dementia. They can have less control over how they feel and express themselves. For instance, someone may overreact to situations, experience abrupt mood swings, or feel agitated. They could also come out as abnormally cold or indifferent [11].

Hence as mentioned in paper [5] music maybe a means of expression for someone with dementia because it has the ability to conveyemotion
As a result of the survey, Figure 1.3-1 depicts that many people think it is important to identify the emotional state of patients before guiding them to necessary music therapy sessions.

Figure 1.3.1 Survey results on the people's knowledge to the approach on music

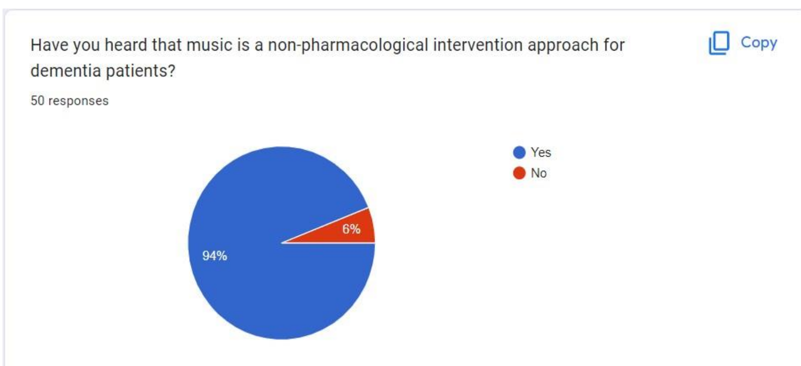
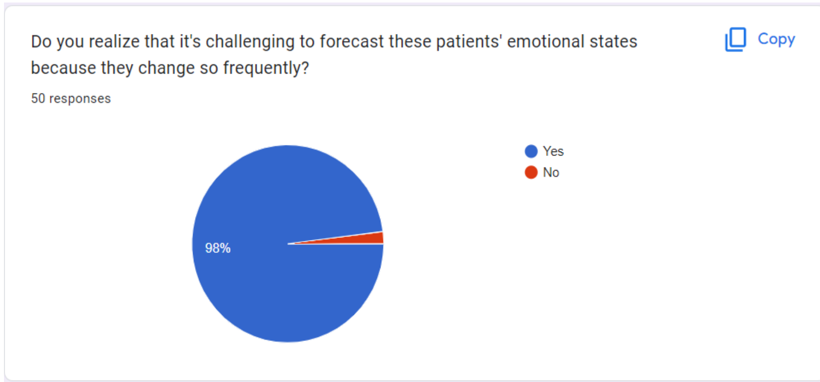


Figure 1.3.2 Survey results on challenges of forecasting the emotional state



Most people with dementia seek the aid of a third party to ease their daily tasks [2]. According to the survey commentary section and Figure 1.3-2, A third party will not be able to effectively identify the patient's emotional condition and direct them to the necessary therapy sessions accordance with those emotions.

According to the survey results, Figure 1.2-1and literature reviews illustrate accessibility issues with the current tools, lack of understanding, difficulty in usage, a limited range of music, and music recommendations thatare ineffective in elevating mood as some other issues.

2. RESEARCH OBJECTIVES

2.1. Main Objective

The goal of assessing a dementia patient's emotional state and making music therapy recommendations in accordance with the emotional state is to enhance that patient's quality of life and lessen behavioral and psychological signs of dementia (BPSD). Agitation, anger, depression, and anxiety are all examples of BPSD, which are prevalent in dementia patients. It has been demonstrated that music therapy helpsthese symptoms by elevating mood, lowering stress, fostering greater social contact, and enhancing general wellbeing. It is intended that a patient's BPSD may be lessened, resulting in an increase in their quality of life overall, by recognizing their emotional state and offering suitable music therapy.

2.2. Specific Objectives

In addition to the main objectives, there are some specific objectives related to the implementation.

- Data Set preparation

The emotional detection algorithm in this project is trained using an existing dataset. The dataset is already labeled with emotional categories after being carefully selected for its relevance and quality, which reduced the requirement for complicated data collection and annotation. For the purpose of ensuring compatibility with the selected machine learning model, data preprocessing, including cleaning and formatting, is carried out. The succeeding stages of model development are accelerated by this streamlined method of dataset preparation.

- Model Training

This phase's main focus is on hyperparameter tuning, training, and model selection. Hyperparameter optimization methods are used to fine-tune an appropriate machine learning model that is selected based on the characteristics of the dataset. For the model to function at its best on the preprocessed dataset, several training iterations are completed. The foundation for the emotional detection system's next stages is formed by this effective model training procedure, which also showed how easily it could be modified to work with current datasets.

- Choosing the appropriate therapeutic music session using an emotion-audio integrationmodule

Provide a music therapy session recommendation module that can reliably identify a patient's emotional condition. Here a database of music therapy sessions is created that have been proven to be successful in treating particular emotional states in dementia patients by doing research and collecting the data. Based on client feedback and new research findings, continuously update and improve the emotion-audio integration module and music therapy session database. To avoid overstimulating or disturbing the patient, make sure the music therapy sessions chosen are appropriate for their cognitive and physical skills.

3. METHODOLOGY

3.1. Methodology

In this research project, we apply a comprehensive approach to improve user experience by making recommendations for music that are driven by emotions. We start by choosing the FER-2013 dataset , a vast collection of facial expressions representing a range of seven emotional states. The training set, the public test set, and the private test set are three subsets that were carefully divided up into this dataset. Our CNN model is trained using the training set, which consists of 28,709 pictures.

We gradually create a CNN model for face expression identification using this dataset. Multiple convolutional layers followed by max-pooling layers make up the CNN design, which enables the model to recognize complex spatial patterns and feature hierarchies in facial pictures. To improve model convergence and feature learning, batch normalization and ReLU activation functions are included in each convolutional layer. Three separate blocks that each gradually decrease spatial dimensions while raising the amount of abstraction make up the CNN's structure. A flatten layer follows these convolutional layers and gets the data ready for fully linked (dense) layers. Further non-linearity and abstraction are introduced by the dense layers. Softmax activation-equipped output layer makes predictions about each image's emotions.

The development of a user-friendly React Native mobile application allows users to quickly and easily capture their facial expressions. The backend server is informed of the associated image URLs while the recorded images are effortlessly transferred to Firebase storage, where they are safely stored. The CNN model is used in the backend, or the brain, of our system to precisely identify emotions from input facial photos. These emotion labels are then conveyed back to the React Native front end via being contained within

JSON requests. The system considers additional contextual information, such as the user's stated age gap, to adjust our approach to the particular preferences and sensitivities of each user. Based on this data, the system uses a sophisticated recommendation technology to offer songs from a personalized music collection that take the user's age gap and the style of music popular in their 20s into account. These musical selections are made in accordance with the identified emotion to provide a unique mood-lifting experience. This approach considers a variety of elements, such as data gathering and preprocessing, model building and training, model integration with the React Native front end, UI design, and performance evaluation. Through this research, we aim to provide valuable insights into the feasibility and effectiveness of personalized emotion-driven music recommendations as a means of enhancing user well-being and satisfaction.

3.1.1. System Architecture

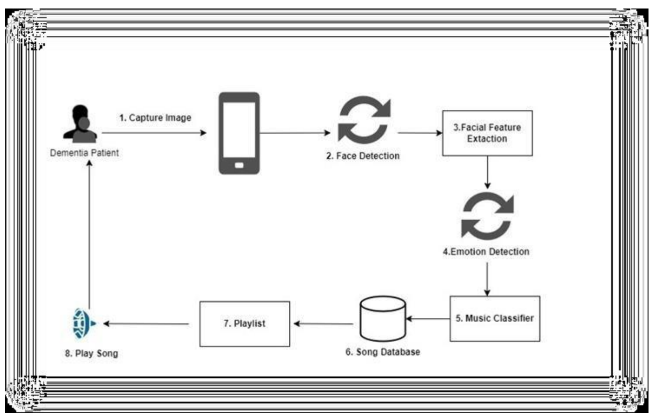


Figure 3.1.1 - Overall System Diagram

Figure 3.1-1 indicates very evident that the backend server manages all aspects of image processing while the mobile application handles UIs, gesture control, and all instructions. While the mobile application has all the client-side features, Figure 3.1-2 depicts a system diagram to describe how the backend of this particular research component functions.

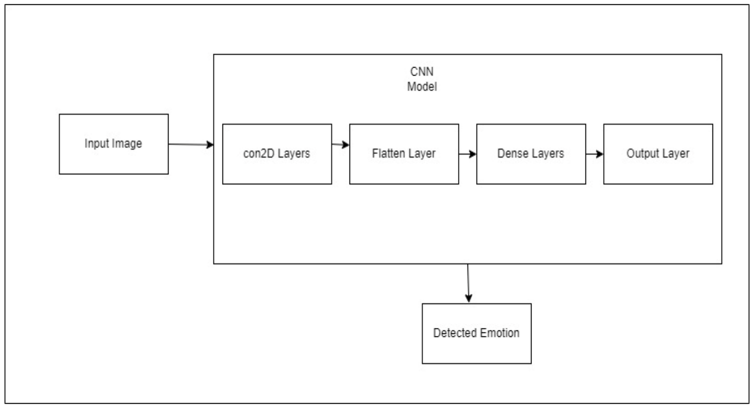


Figure 3.1.2 - System Diagram for the Individual Component

3.1.2. Data collection methods

The key decision to use the Kaggle FER-2013 dataset for this investigation was motivated by its unique features and applicability to the study's goals. FER-2013 is a thorough and freely accessible database of facial expressions that includes seven different emotional states: anger, disgust, fear, happiness, sorrow, surprise, and neutral. This dataset provides a wide range of emotions, which makes it ideally suited for developing and testing our emotion recognition model.

Tools

- React Native
    - To implement the mobile application
- Kaggle Notebook
    - To train the machine learning models
- PyCharm
    - To implement the final backend
- Postman
    - To test the mobile application

Technologies

- python
- Keras
- Fast API

  To implement the backend API for the mobile application
- jsx
  - Used to implement mobile application with react native
- TensorFlow
  - Used to implement the model

Table 3.3.1 - External tools

Description	**Tools**
Version Controlling	Gitlab
Team connectivity	Teams, WhatsApp

3.2. Commercialization aspects of the product

In our pursuit of commercializing our mobile application, which holds immense potential to benefit individuals worldwide, we recognize the need for a multifaceted approach that aligns with current global trends and user behaviors. Given that our application operates in the international language of English, its reach extends far beyond the borders of Sri Lanka, making it accessible and relevant to a global audience. To ensure swift and widespread adoption, we have strategically identified several key avenues for commercialization, with a primary focus on harnessing the immense power of social media platforms.

In today's digital landscape, social media stands out as a dynamic and influential force. With millions of users spending a substantial portion of their daily lives on platforms such as Facebook, WhatsApp, Instagram, and YouTube, leveraging these channels for advertising and promotion is not only prudent but essential. Through targeted and engaging adverts on these platforms, we can effectively introduce our application to a vast and diverse consumer base, transcending geographical boundaries.

Furthermore, we have recognized the importance of content providers in shaping trends and boosting user engagement. Platforms like YouTube, Twitch, and Trovo are home to a growing community of content creators with large followings. By selectively sponsoring and supporting these influencers, we can use their reach and influence to raise awareness about our system within their respective communities. This spontaneous recommendation can help us create confidence and credibility for our application.

To broaden our reach, we intend to work closely with healthcare organizations such as hospitals and clinics. Within their particular communities, these institutions provide reliable sources of knowledge and care. We can design targeted awareness programs that appeal to both healthcare professionals and patients by collaborating with them. This strategy ensures that our application reaches people from all walks of life, regardless of financial status.

We recognize the lasting power of conventional media in addition to digital platforms. For example, leaflets can be a useful instrument for increasing public awareness of our product, particularly within rural communities. We also recognize the importance of radio and podcasts, which continue to have a loyal following. Sponsoring radio shows and podcasts that cater to our target audience can be a beneficial outlet for advertising our product and its benefits.

3.3. Testing and Implementation

3.3.1. Implementation

To develop an app that utilizes emotion recognition to play music, we need to gather relevant information and identify a suitable dataset to train the model. The focus will be on interpreting emotions and selecting appropriate music based on the analysis.
Before starting the implementation, first a requirement analysis must be done. Therefore, user requirements functional requirements and non-functional requirements were gathered as below.

Functional requirements

- Extract features from the image
- Identify the objects of the image
- Describing the colors of the image
- Describe the main features of the image
- Generate meaningful captions
- Further description of the image using surrounded objects near mainsubject

Non-functional requirements

- Usability
- Accuracy
- Availability
- Well optimized for cloud/mobile use

User requirements

- User should have a mobile phone to use the application
- User should have an English knowledge to understand the guidelines
- User should be able to hear
- User should have a simple knowledge to use a mobile application

For the implementation in the image interpretation, it has both server side and client-side implementations, client-side implementations mean the mobile application and the server-side implementation means the image processing process. In the mobile application implementation, it is developed using react native framework.

```python
#dense 1
model.add(Dense(2*2*2*num_features))
model.add(BatchNormalization())
model.add(Activation('relu'))

#dense 2
model.add(Dense(2*2*num_features))
model.add(BatchNormalization())
model.add(Activation('relu'))

#dense 3
model.add(Dense(2*num_features))
model.add(BatchNormalization())
model.add(Activation('relu'))

#output layer
model.add(Dense(num_classes, activation='softmax'))

model.compile(loss='categorical_crossentropy',
              optimizer=Adam(lr=0.001, beta_1=0.9, beta_2=0.999, epsilon=1e-7),
              metrics=['accuracy'])

model.summary()
```

```python
model = Sequential()

#1
model.add(Conv2D(2*2*num_features, kernel_size=(3, 3), input_shape=(width, height, 1), data_format='channels_last'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(Conv2D(2*2*num_features, kernel_size=(3, 3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2), strides=(2, 2)))

#2
model.add(Conv2D(2*num_features, kernel_size=(3, 3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(Conv2D(2*num_features, kernel_size=(3, 3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2), strides=(2, 2)))

#3
model.add(Conv2D(num_features, kernel_size=(3, 3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(Conv2D(num_features, kernel_size=(3, 3), padding='same'))
model.add(BatchNormalization())
model.add(Activation('relu'))
model.add(MaxPooling2D(pool_size=(2, 2), strides=(2, 2)))

#flatten
model.add(Flatten())
```

Figure 3.3.1 – Usage of the CNN model

The CNN model that is being used is carefully crafted to extract complex information from facial photos, enabling the recognition of subtle emotional expressions. Multiple convolutional layers with batch normalization and Rectified Linear Unit (ReLU) activation functions are the first layers in the model. The learnable filters are progressively slid across the input images by these neural layers, catching crucial details like edges and face elements. Following, max-pooling layers downscale feature maps while maintaining essential data. The output is reshaped by the flattening layer to make room for fully connected dense layers. Higher-level features are gradually abstracted from the retrieved representations by these dense layers. The output layer uses softmax activation to compute probabilities and has seven neurons that correspond to the different emotion classes. The convolutional layers serve as feature extractors throughout this process, learning to recognize important patterns and shapes, such as the curve of a grin or the furrow of brows, indicative of various emotions. Through backpropagation and Adam optimization, the model is trained to minimize a loss function while optimizing weights at a particular learning rate. After that, the model's effectiveness is assessed on validation and test datasets to make sure it has the ability to generalize and correctly identify emotions in facial photos that cannot be seen.

```
data_generator = ImageDataGenerator(
                        featurewise_center=False,
                        featurewise_std_normalization=False,
                        rotation_range=10,
                        width_shift_range=0.1,
                        height_shift_range=0.1,
                        zoom_range=.1,
                        horizontal_flip=True)


es = EarlyStopping(monitor='val_loss', patience = 10, mode = 'min', restore_best_weights=True)

history = model.fit_generator(data_generator.flow(train_X, train_Y, batch_size),
                                steps_per_epoch=len(train_X) / batch_size,
                                epochs=num_epochs,
                                verbose=2,
                                callbacks = [es],
                                validation_data=(val_X, val_Y))
```

Figure 3.3.2 – Training of the proposed model

A key element of this research is the training phase of the emotion detection model, which is distinguished by a purposeful effort to increase both the model's robustness and its capacity for successful generalization. The training set, which consists of 28,709 images, is the devoted subset of the FER-2013 dataset on which the model is thoroughly trained. A data augmentation approach is thoroughly implemented to support the model's resilience and guarantee its adaptability to various facial expressions. To provide variety to the training data, the ImageDataGenerator, a flexible tool, is used. This augmentation method expands the dataset's breadth and exposes the model to a wider range of facial emotions and changes by performing random rotations, width and height shifts, zooming, and horizontal flipping.

An early halting mechanism is intentionally used to prevent overfitting, a serious risk in DL. Throughout training, EarlyStopping supervises the model's progress by continuously keeping an eye on its performance on a different validation dataset. The validation loss is closely monitored, and if no improvement is seen after a predetermined number of training epochs, which is 10 in this case, the training process is politely terminated. By avoiding overfitting to the training data, this keeps the model's capacity to generalize to novel, unseen face expressions intact. This strategy also guarantees the model performs at its best by returning it to its ideal state, providing a strong answer for precise emotion recognition. Essentially, the combination of early stopping and data augmentation creates a model that is capable of distinguishing complex facial features while reducing the risk of overfitting, ultimately acting as a strong part of our emotion-driven music recommendation system.

The journey in our user-centric system starts with users uploading emotional images via the 'uploadEmotionImage' feature. The URI, filename, and type of the picture are parameters for this function, which is managed on the front end. Giving each image a distinct URL, it smoothly delivers this image data to Firebase Cloud Storage for safe storage and retrieval. This URL is subsequently sent to the backend server, when the 'predictEmotion' function assumes center stage. A post request is orchestrated to the server's '/predict_emotion' endpoint by the 'predictEmotion' function, which is enabled by Axios. The emotion detecting process is started at this crucial stage. Behind the scenes, a pre-trained CNN model carefully examines the facial image and accurately identifies the prevalent emotion. The front end receives an immediate JSON response that contains the emotion that was identified.

```
export const fetchAudio = async () => {
  let audios = [];
  const emotion = await getFromStorage("CURRENT_EMOTION");
  console.log(`current emotion is ${emotion}`);
  const age = await getFromStorage("age");
  const ageRange = ageRangeGetter(age);

  const q = query(
    collection(db, "audio"),
    where("category", "==", emotion),
    where("ageRange", "==", ageRange)
  );
  await getDocs(q)
    .then((snapshot) => {
      snapshot.docs.map((doc) => {
        const track = new SongModel(
          doc.data()["url"],
          doc.data()["artist"],
          doc.data()["category"],
          doc.data()["ageRange"]
        );
        audios.push(track);
      });
    })
    .catch((err) => {
      console.log(`Error in get auidos ${err}`);
      return [];
    });
  return audios;
};
```

Figure 3.3.3 - Getting the person age and the predicted emotion

Following the identification and localization of the user's current feeling, our system sets out on a harmonic musical journey to enhance the emotional experience. This procedure is expertly and carefully orchestrated by the 'fetchAudio' function, which takes center stage. In order to provide a customized musical experience, it begins by obtaining the user's current feeling. The user's age is also considered, which is an important consideration. A selection of music that is age-appropriate is made possible by using age information to identify the best age range. The function creates a focused query to the music database using these inputs. It searches through the collection of audio tracks using Firebase's Firestore to find ones that fit the age range and sentiment that have been identified. The outcome is a carefully curated collection of audio tracks, each one carefully chosen to relate to the user's emotional state and age range. These songs capture a wide variety of moods and genres and are ready to take the listener on a relaxing audio journey. The 'fetchAudio' function essentially acts as a link between emotion recognition and music selection. The user's emotional context is flawlessly matched with every musical note and rhythm, increasing the emotional experience and producing a completely immersive and unique music selection. To keep the user engaged with the application a slideshow will be presented with the music screen as well.

The feasibility studies listed below are carried out during the development process for the implementation feasibility study. The following description provides schedule feasibility, technical feasibility, and economic feasibility elements.

Schedule Feasibility:

The suggested system must be finished in the specified period of time. Each step should have a deadline to guarantee a high-quality final output. The time constraints for each work will be displayed on the Gantt chart.

Technical Feasibility:

Members of the research team should have some familiarity with machine learning techniques and mobile application development technologies. All members of the research team should be familiar with the computer programming languages needed to implement the suggested application.

Economy feasibility:

Cost limitations for the product's resources are required. All participants should fall inside the budget. The strategy ought to be more thorough and less expensive.

3.3.2. Testing

The product will be tested using a range of testing approaches, such as unit testing, integration testing, and user acceptability testing, because appropriate testing guarantees that faults and bugs are detected early in the life cycle of the application.

As soon as all of the testing is finished, it should be made available. Prior to the product's release, any issues that surfaced during the testing phase should be fixed. After user testing is finished and approved, this mobile-based application for dementia sufferers will be made available.

Some of the test cases used to test the product are included below, along with screenshots.

Table 3.3.2 - Test cases for Emotion generation and music player

Test Case #	Test case	Result
001	Camera opens from the open camera button	Pass
002	All the buttons and widgets are visible	Pass
003	Navigate for pages through buttons	Pass
004	Vibration works when touch the buttons	Pass
005	Image successfully captured after opening the camera	Pass
006	Image successfully uploaded for the algorithms	Pass
007	Detect the emotion	Pass
008	Detected emotion is displayed in the screen	Pass
009	Listen to music button is clicked	Pass
010	Appropriate Music is generated to the user	Pass
011	Slideshow is working properly	Pass

Above table shows the test cases done for emotion detection and generation of the music player. Also, this application was tested by using two different end users and table 3.3.2 shows the results of it. For that we contacted one person having mild dementia and another patient who has moderate dementia. Those people are respectively named in the table as User 1 and User 2.

Table 3.3.3 - Test cases done by end users

Test Case #	**Test case**	**User 1**	**User 2**
001	Open the application without any error	Opened the application without any issue	Opened the application without any issue
002	Camera opens from the Music icon in the home page.	Able to open the camera	Able to open the camera
003	Navigate for pages through buttons	Navigated through all the pages	Navigated through all the pages
004	Vibration works when navigating through buttons	Navigated through the pages with the help of vibration	Navigated through all the pages
005	Able to capture the image successfully	Captured the image without any issue	Captured the image without any issue

Finally, we checked the applications on various Android devices and OS versions to wrap up the testing process. The test cases for it are displayed in table 3.3.3 below.

Table 3.3.4 - Test cases for devises with different OS

Test Case #	Device	OS	Version issues	Issues with the interfaces
001	Xiaomi 11T	Android 13	No issues	No issues
002	Galaxy M14 5G	Android 11	No issues	No issues
003	Samsung galaxy S8	Android 9	No issues	No issues

4. RESULTS AND DISCUSSION

4.1. Results

The system for captioning images provides the person's present emotional state for nearly every image. As the features of the detected image is extracted with this CNN model the accuracy of the generating emotions is also high as shown in figure 4.1.1 which is a confusion matrix generated to the model implemented.

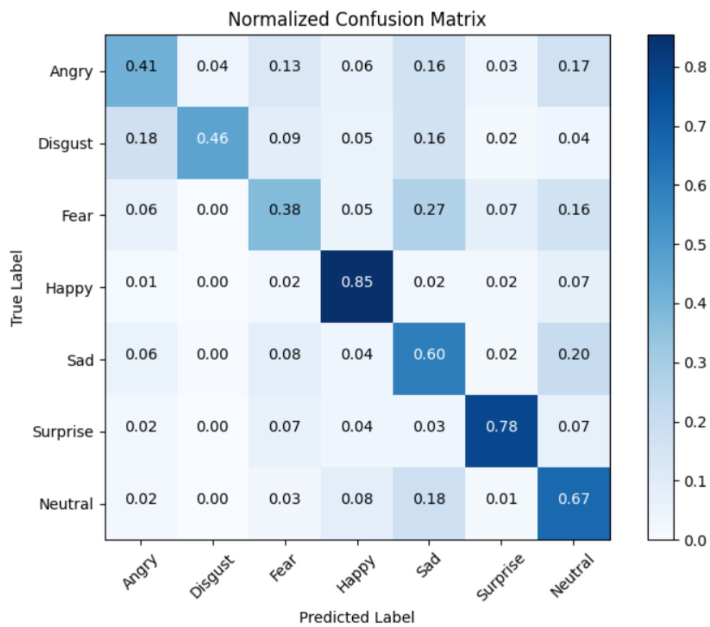


Figure 4.1.1 - Accuracy of the generated attributes

As shown above, the accuracy for the attributes happy, sad, surprise, neutral, fear, disgust and angry are respectively 88%,60%,78%,67%,38%,46%,41%.

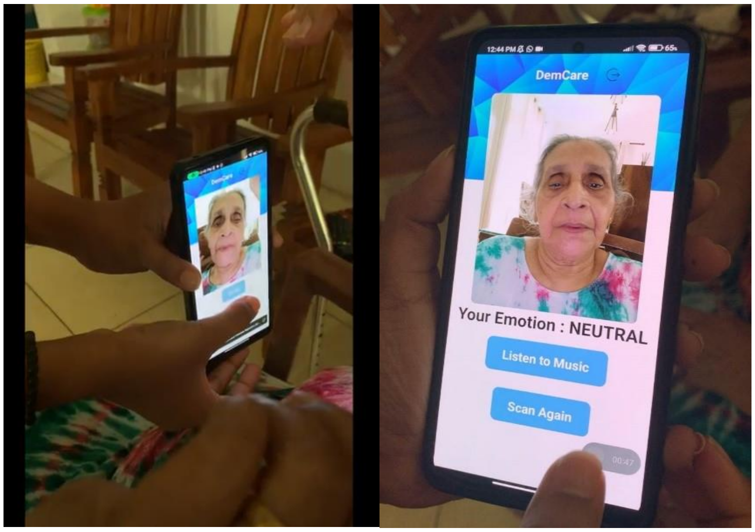


Figure 4.1.2 - Sample image tested with a moderate dementia patient and the detected emotion

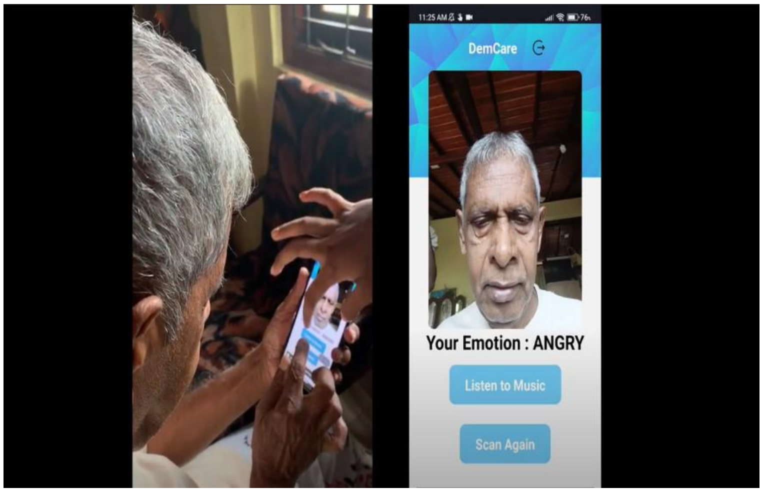


Figure 4.1.3 – Sample image of a mild dementia individual and the result

Figure 4.1.2 and Figure 4.1.3 displays two outcomes that were gathered from individuals. As you can see, the projected emotions closely match the patient's actual emotional state at the moment. Thus, it is evident that the mobile application's implementation provides highly accurate forecasts.

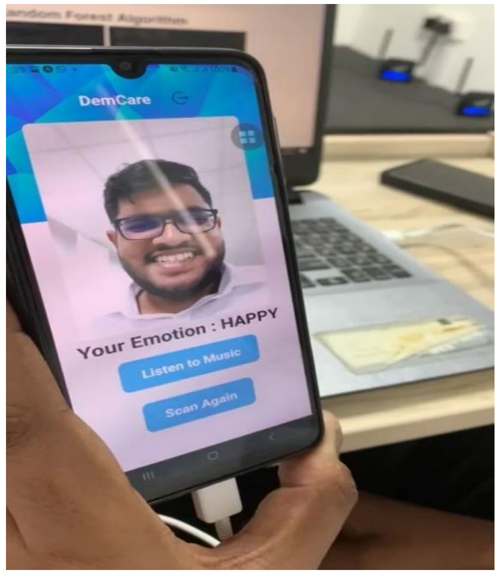


Figure 4.1.4 - Sample image tested with a team member

4.2. Research Findings

The study's findings highlight the usefulness and efficacy of the suggested emotion-driven music recommendation system. The system recognized a variety of emotions, such as happy, sadness, anger, and neutral, with respectable accuracy by using a pre-trained CNN model for emotion recognition. The model's performance was improved and overfitting was minimized by the use of data augmentation approaches, guaranteeing that it could successfully generalize to new data. The personalized "musicService" collection showcased its skill at selecting musical pieces that connected with the emotions and age range of consumers, enhancing their mood and emotional experience. User engagement was greatly increased by the user-centric strategy, which included seamless image capture, emotion detection, and music selection. The architecture of the system, which uses Firebase for image storage and retrieval and Firestore for music database queries, also proved to be effective and scalable. Age was taken into consideration when choosing music, ensuring that choices were made that catered to a variety of user demographics.

These research findings show how the system can be used in the real world to enhance mental health, provide entertainment, and improve user experiences through emotional involvement.

4.3. Discussion

Throughout the research's development and evaluation phases, a number of important factors were found. In the past, some research on the topic was conducted to gain a better understanding of these dementia patients' behaviors. We conducted some preliminary research and looked for some first-hand knowledge. We then made the decision to run an online survey in order to gather some important research data. The survey was conducted by sending out a Google form.

The system first creates the appropriate emotion based on the provided image using a CNN algorithm. Here, the image's URL is sent to the backend server, where the CNN model we developed is used to make the prediction. The predicted emotion is subsequently delivered as a JSON request to the front end. A customized music playlist will be created from a pre-made music library using this emotion and the user's registered age with a slideshow in the UI to keep the user engaged. I was able to do every task and get a high accuracy for the prediction, as suggested in the proposal. There are a few things that should be implemented during the testing procedure. When taken with a smartphone, the image quality degrades. Thus, it has an impact on the output's accuracy. The accuracy can be improved further than the current findings if we can capture the image with a high-quality image.

However, the current image capture result produces considerable output that the user can understand. The user will initially capture the face from the mobile device. The backend will then take in the input and gather it for processing. The CNN architecture will process the image through numerous layers in the backend before predicting the emotion.

Additionally, the system's user-friendly, minimalistic design will help the user navigate their way through the process.

4.4. Summary of student contribution

Student: Hiththatiyage D.K– IT20162696

Research component: Implementing an emotion-based music player

Task: This component is aligned with to develop a system to interpret the emotion of a captured face. The inputs of this component are the image of the individual and the registered age, and the output of the predicted emotion will be displayed to the user in a printed format. Furthermore, the minimalistic UI of the application will help the patient when navigating through the application.

Tasks completed:

- Developed the mobile application
- Implemented the backend using Fast-API to get the image interpretation services by the application
- Developed image captioning and emotion detection model.
- Implemented the algorithm to extract the image content from the image
- Developed haptic feedback system in the proposed mobile application.

5. CONCLUSION

In this study, numerous DL and image processing techniques are used to build a unique image processing approach. When compared to state-of-the-art methods, the developed methodology performed noticeably better on the FER-2013 dataset. This method may be included into hardware platforms like smartphones, reducing the difficulties that people with dementia deal with on a regular basis. This strategy has limitations because the capability of the smartphone camera mostly determines how accurate the overall system is.. A better external camera could be incorporated into the remedy to prevent this. This method of image captioning can be improved in the future by following the aforementioned advice to increase accuracy.

Additionally, as the music library was hand created, it has some level limitations. To make the experience more seamless, it would be preferable to leverage an existing music collection. Alternatively, a large collection must be created.

REFERENCES

[1] G. K. -H. Pang and E. Kwong, "Considerations and design on apps for elderly with mild-to-moderate dementia," *2015 International Conference on Information Networking (ICOIN)*, Cambodia, 2015, pp. 348-353, doi: 10.1109/ICOIN.2015.7057910.

[2] Sposaro F, Danielson J, Tyson G. iWander: An Android application for dementia patients. Annu Int Conf IEEE Eng Med Biol Soc. 2010;2010:3875-8. doi: 10.1109/IEMBS.2010.5627669. PMID: 21097072.

[3] Philippa Riley, Norman Alm, Alan Newell,An interactive tool to promote musical creativity in people with dementia, Computers in Human Behavior, Volume 25, Issue 3,2009,Pages 599-608,ISSN 0747-5632, https://doi.org/10.1016/j.chb.2008.08.014.

[4] Davi de Oliveira Cruz, Carlos Chechetti, Sonia Maria Dozzi Brucki, Leonel Tadao Takada, Fátima L.S. Nunes,A comprehensive systematic review on mobile applications to support dementia patients,Pervasive and Mobile Computing,Volume 90,2023,101757,ISSN 1574-1192,

[5] Emotion Based Music Player
Hafeez Kabani1, Sharik Khan2, Omar Khan3, Shabana Tadvi4

[6] Cunningham, S. et al. (2019) "Assessing wellbeing in people living with dementia using reminiscence music with a mobile app (memory tracks): A mixed methods cohort study," Journal of Healthcare Engineering, 2019, pp. 1–10. Available at: https://doi.org/10.1155/2019/8924273.

[7] Nezerwa, Martine & Wright, Robert & Howansky, Stefan & Terranova, Jake & Carlsson, Xavier & Robb, John & Coppola, Jean. (2014). Alive Inside: Developing mobile apps for the cognitively impaired. 1-5. 10.1109/LISAT.2014.6845228.

[8] Sourina, Olga & Liu, Yisi & Nguyen, Minh Khoa. (2011). Real-time EEG-based emotion recognition for music therapy. Journal on Multimodal User Interfaces. 5. 10.1007/s12193-011-0080-6.

[9] Athavle, Madhuri. (2021). Music Recommendation Based on Face Emotion Recognition. Journal of Informatics Electrical and Electronics Engineering (JIEEE). 2. 1-11. 10.54060/JIEEE/002.02.018.

[10] van der Steen JT, van Soest-Poortvliet MC, van der Wouden JC, Bruinsma MS, Scholten RJ, Vink AC. Music-based therapeutic interventions for people with dementia. *Cochrane Database Syst Rev*. 2017;5(5):CD003477. Published 2017 May 2. doi:10.1002/14651858.CD003477.pub3
.

[11] Athavle, Madhuri. (2021). Music Recommendation Based on Face Emotion Recognition. Journal of Informatics Electrical and Electronics Engineering (JIEEE). 2. 1-11. 10.54060/JIEEE/002.02.018.
The psychological and emotional impact of dementia (2019) Alzheimer's Society. Available at: https://www.alzheimers.org.uk/get-support/help-dementia-care/understanding-supporting-person-dementia-psychological-emotional-impact#:~:text=People%20with%20dementia%20often%20experience,distant%20or%20uninterested%20in%20things. (Accessed: March 20, 2023).

APPENDICES

Appendix A: Online Survey

Survey conducted in order to get details on developing a mobile app for dementia patients.

Dear Respondent,
I'm a Final year student from the department of computer science and software engineering, faculty of computing , SLIIT

I'm researching dementia patients to identify their emotions and create music therapy sessions for them in accordance with those emotions.. This survey is conducted to gather some date required to proceed with the research.

udzhiththatiyage@gmail.com (not shared) Switch account

* Required

Gender *

- Female
- Male
- Other:

Age *

Age *

- ○ Below 18
- ○ 18-25
- ○ 26-30
- ○ 31-40
- ○ 41-60
- ○ Above 60

Have you heard about Dementia *

- ○ Yes
- ○ No

Have you ever associated a Dementia Patient *

- ○ Yes
- ○ No
- ○ Maybe

How often do you listen to music? *

- ◯ Everyday
- ◯ Few times a week
- ◯ Once a Week
- ◯ A few times a month
- ◯ Rarely or never

Have you heard that music is a non-pharmacological intervention approach for dementia patients? *

- ◯ Yes
- ◯ No

Do you realize that it's challenging to forecast these patients' emotional states because they change so frequently? *

- ◯ Yes
- ◯ No

What ideas do you have for introducing music therapy sessions into patient care? *

- [ ] Getting help from a third party
- [ ] Using an assistive tool
- [ ] Using a music player
- [ ] Using a tool with tactile feedback system

Do you believe getting assistance from a third party to determine the mood and direct them to the appropriate music is effective? *

- ○ Yes
- ○ No

Do you believe a digital assistant is required to anticipate patients' emotional states and suggest music to them based on those moods? *

- ○ Yes
- ○ No

Have you heard of any such tools existing dementia patients? *

- ◯ Yes
- ◯ No

If so, What do you think of the accuracy of those available tools?

- ☐ very accurate
- ☐ Somewhat accurate
- ☐ Not very accurate
- ☐ Not accurate at all
- ☐ Not sure/don't know

What are the flows you see in the available tools?

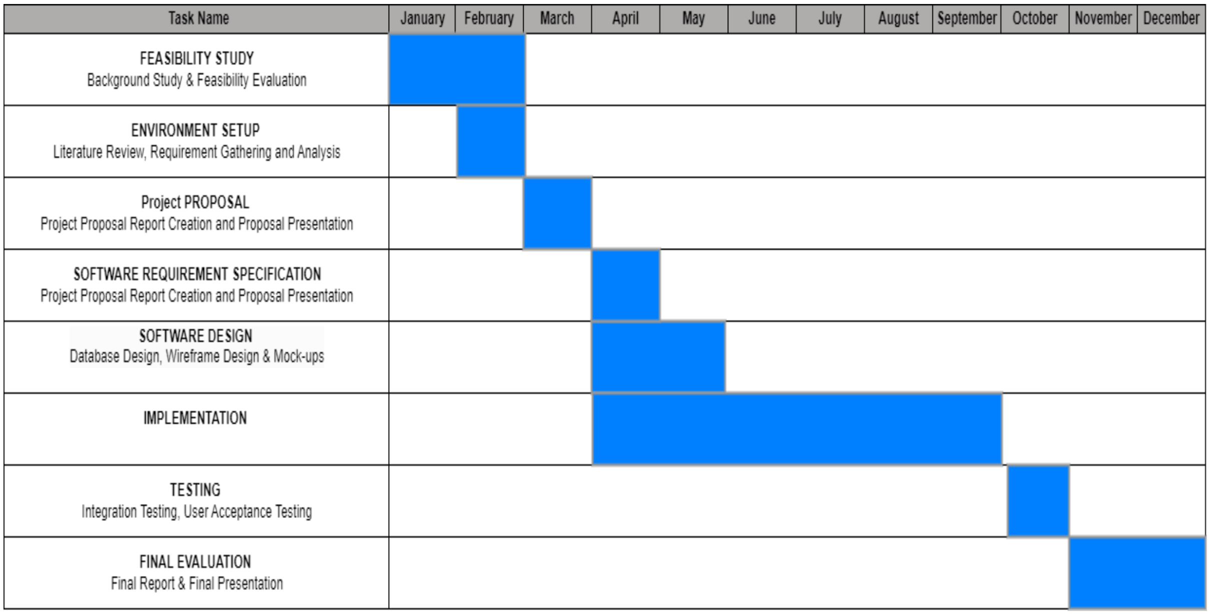
- ☐ Not user friendly for dementia patients
- ☐ Emotion detection is not accurate
- ☐ Difficult to use
- ☐ Limited selection of music
- ☐ Music recommendations are not effective in improving themood

What suggestions do you have for making our tool better when it comes to focusing on analyzing a patient's emotions and playing music in accordance with that? *

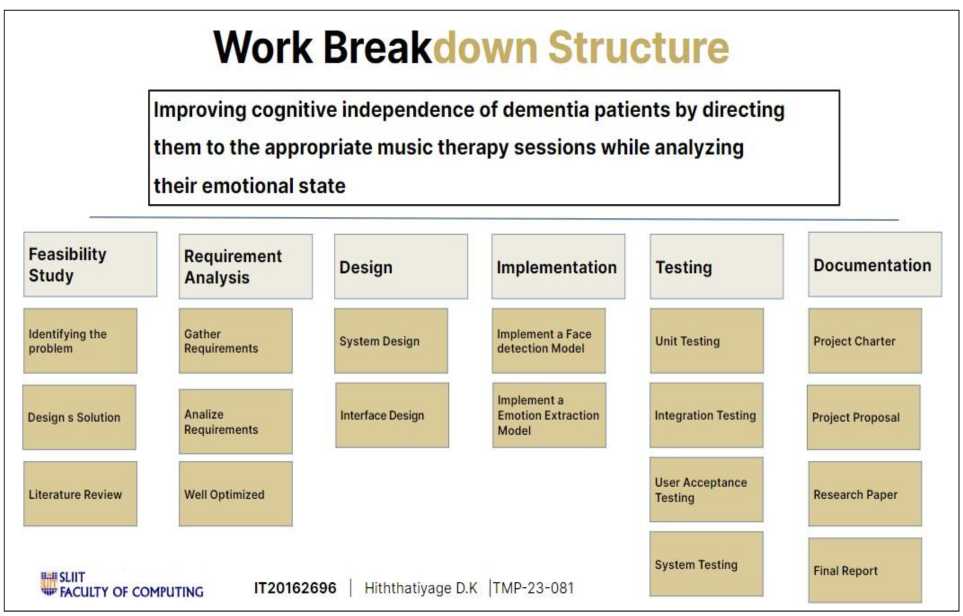
- [ ] Improving the accuracy of emotion detection
- [ ] Providing more diverse music options
- [ ] Use simple English and simple UIs
- [ ] Making the tool more user-friendly and accessible for patients with varying levels of cognitive ability.
- [ ] Personalizing the music recommendations based on individual age

Submit   Clear form

Appendix B: Gantt chart



Appendix C: Work Breakdown Structure



Appendix D: Plagiarism report

IT20162696_Individual_Thesis

ORIGINALITY REPORT

3%	2%	0%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Luton Sixth Form College Student Paper	1 %
2	Submitted to University of the Philippines Los Banos Student Paper	<1 %
3	Submitted to University of Suffolk Student Paper	<1 %
4	Submitted to Harare Institute of Technology Student Paper	<1 %
5	Submitted to TAFE Queensland Brisbane Student Paper	<1 %
6	www.researchgate.net Internet Source	<1 %
7	www.dcnicn.com Internet Source	<1 %
8	alta2016.alta.asn.au Internet Source	<1 %